Blood Donors Studies BioResource
Data Access Policy

Table of Contents

1. Overview of the Blood Donors Studies BioResource

a. The Blood Donors Studies BioResource (BDSB) will include de-personalised genetic, biological and lifestyle information linked to health outcomes for up to 325,000 blood donors recruited into the INTERVAL, COMPARE and STRIDES BioResource studies. Collectively these studies are referred to as the Blood Donors Studies and the resulting collection of data and biological samples forms the Blood Donors Studies BioResource. The studies and the BDSB are managed through the NIHR Blood and Transplant Research Unit (BTRU) in Donor Health and Genomics.

b. Participants in the INTERVAL, COMPARE and STRIDES BioResource studies agreed for their de-personalised data to be used for health-related research purposes. The BDSB will support population health and biomedical research by providing access to pseudonymised data from this large and richly characterised cohort of healthy volunteers. This is expected to lead to a better understanding of the determinants of important diseases and therefore contribute to improvements in risk assessment and treatment.

c. A central feature of the BDSB is for pseudonymised datasets to be systematically accessible to *bona fide* researchers in the wider scientific community. This includes those working in universities, charities, government agencies or commercial companies in the UK and abroad.

d. The release of data is regulated by the BDSB Data Access Committee (DAC) which includes senior members from the organisations involved in the studies: University of Cambridge, NHS Blood and Transplant (NHSBT), the Wellcome Sanger Institute and the University of Oxford.

e. The DAC will review the project's scientific excellence and alignment of the proposal with the overall aims of the database, the research team's experience and capability to conduct the proposed study, and the suitability of the data and any risk to participant confidentiality.

f. The data that is available in the BDSB Research Database is listed in the Appendix and the Study Data Dictionaries.

2. Objective of the Data Access Policy

a. The objective of the Data Access Policy is to facilitate access to the data in the BDSB so that the data get the widest possible usage, whilst ensuring that such access and usage is consistent with the consent given by the blood donors participating in the INTERVAL, COMPARE and STRIDES BioResource studies and the wider public interest.

b. Applicants who apply to use the data from the BioResource will be required to explain how their research project enables detailed study of the health of blood donors and broader public health and biomedical issues.

c. Applicants can only request the specific data required to fulfil the aims of the project (e.g. they cannot submit a broad project and request all of the data).

d. At all times, decisions to grant access should consider the commitments given to the blood donors when they consented to take part: (i) to ensure that any uses of the BioResource are consistent with its stated aims; (ii) to protect participants' confidentiality; (iii) to ensure that research projects have relevant scientific approval if required, and (iv) to make information publicly available about the uses of the BioResource.

e. It is intended that procedures in the Data Access Policy are clear and transparent and are implemented in a manner which is proportionate, accountable and fair.

f. The procedures provide a framework for addressing and determining access issues. They deliberately do not prescribe what will be done in each and every circumstance because the BDSB DAC cannot predict the nature of access requests that it will receive over the long-term. Built into the procedures is sufficient flexibility to address both the expected and the unexpected, including the ability to revise them in the light of practical experience.

g. The functioning of the Data Access Policy will be kept under review by the BDSB DAC. Input from blood donors, researchers, funders, and other interested parties will be taken into account. Any unresolved issues will be discussed with the Blood Donors Studies Steering Committee.

h. The DAC will amend the Data Access Policy as required on a periodic basis.

3. Factors Affecting Access

3.1 Managed Access Approach

a. The BDSB aims to provide fair, consistent and transparent access to the database in order to promote health-related research by *bona fide* researchers that is in the public interest. No distinction will be made between applicants on the basis of whether they are:
   - From the UK or abroad; and/or
   - Charitable, academic, governmental or commercial entities.

b. However due to limitations on resources, priority will be given to applicants whose work is closely aligned to the primary aims of the studies (i.e. enhancing the safety of blood donation and ensuring sustainability of the blood supply) and to strategic collaborations involving the study principal investigators.

c. The BDSB will apply a standard set of criteria (subject to on-going review and amendment by the BDSB DAC) to the assessment of all applications, including the compatibility of the research project with the purposes of the BioResource; the feasibility of the research project; and the arrangements for managing data.

d. The BDSB wishes to encourage collaboration between prospective researchers in order to increase the efficient use of the data. For example, when proposed research involves overlapping uses of the data, the BioResource team may initiate contact (in confidence and with mutual consent) between applicants.

e. Although collaboration is encouraged, there will be no restrictions on the number of applicants who can be provided with the same or overlapping data that are contained within the BioResource.

f. This approach is intended to encourage rapid reporting of findings and different approaches to the analysis and interpretation of the data, as well as allowing researchers to confirm or refute published findings based on the BioResource (i.e. representing a robust form of peer review that is consistent with the aim of facilitating use by all *bona fide* researchers).

3.2 Legal and Ethics Approval

A favourable opinion will be sought from a Research Ethics Committee (REC) to include generic ethical approval for research projects using de-personalised extracts from the BDSB, under conditions agreed with the REC, without requirement for researchers to apply individually to the REC for approval.

3.3 Intellectual Property Rights

The University of Cambridge will have no claim over any inventions that are developed by applicant(s) using the BioResource (unless they are used to restrict health-related research or access to healthcare unreasonably).

3.4 Restricted Data

Currently, access to datasets provided by NHS Digital is restricted to employees of the University of Cambridge as per the terms of the Data Sharing Agreement. Therefore these datasets will not be shared with external researchers, until an amended agreement which allows this is in place. Similar restrictions may be applied to other datasets provided by third parties.

3.5 Cost

The Applicant(s) may be asked to pay (on a cost-recovery basis) a variable charge depending on the data that is requested for the research project.

4. Application and Review Process

4.1 Application for Access to Data

a. Researchers requesting access to the data are required to make a formal application.

b. There are 4 stages to the application process:
   i. Registration
   ii. Preliminary application
   iii. Application
   iv. Agreements

c. An overview of the process and the interactions between the different parties involved in the application process is shown in the diagram below.

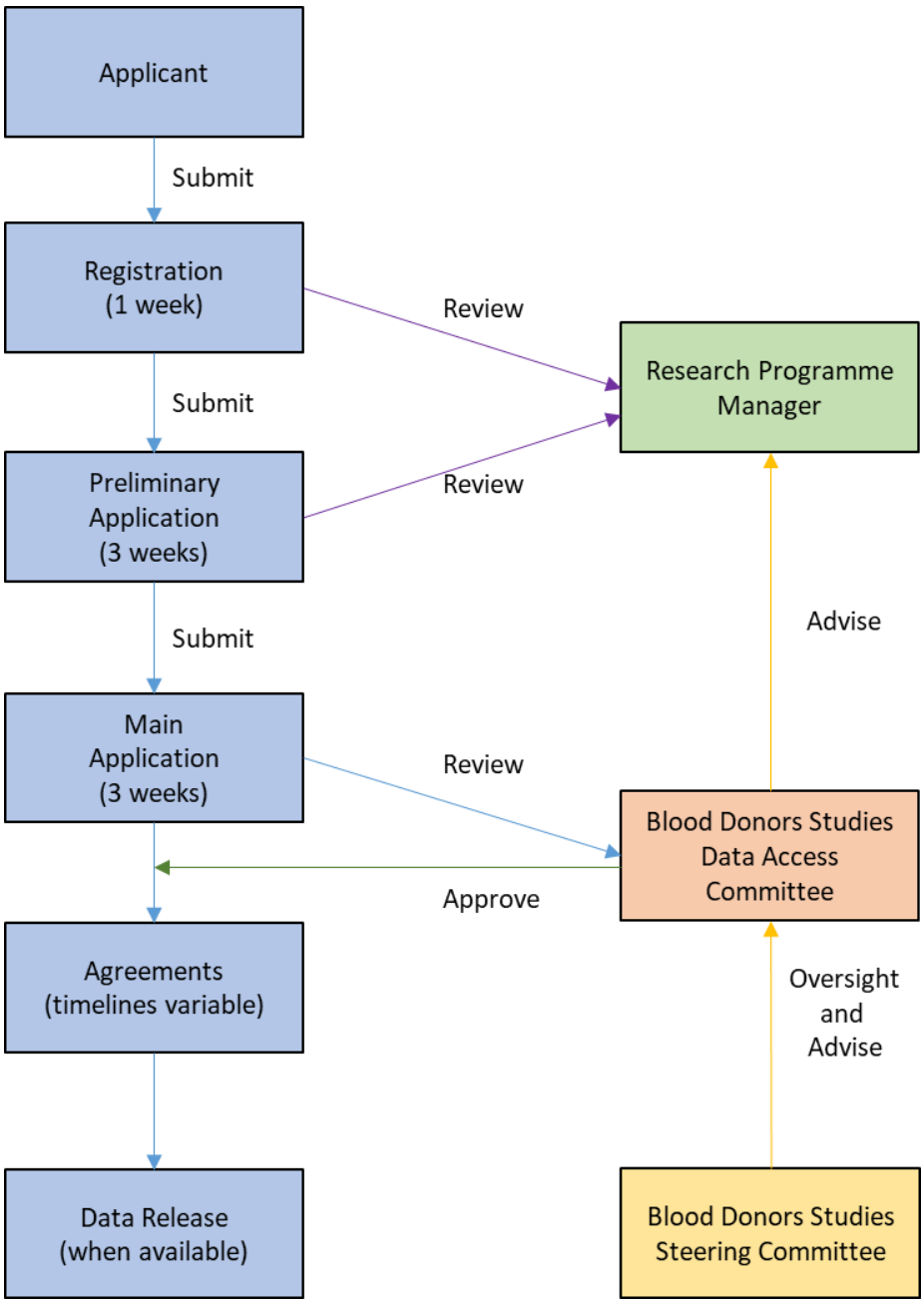

Figure 1: An overview of the process and the interactions between the different parties involved in the data application process.

4.2 Roles and responsibilities

a. **BTRU Research Programme Manager**. The Research Programme Manager is responsible for the smooth running of the access process. They:
- Respond to enquiries from prospective applicants;
- Receive registrations and confirm applicants' *bona fides*;
- Review and respond to preliminary applications;
- Co-ordinate review of full applications by the DAC;

- Manage agreements and provision of data for approved applications;
- Track study outputs and ensure data is destroyed at the end of the study;
- Maintain a database of applications and outputs from the BioResource.

b. **Data Access Committee (DAC)**. The DAC has overall responsibility for the access procedures and all access decisions. The committee reviews full applications and provides advice as required to the BTRU Research Programme Manager. It approves the access policy and may make changes to it if necessary.

c. The Committee includes individuals with expertise in the use of genetic-epidemiological research resources. It includes the Principal Investigators of the Blood Donors Studies, collaborators at the University of Oxford, the Wellcome Sanger Institute and NHSBT and blood donor representatives (lay members of the public).

d. The role of the blood donor representatives is to provide advice on whether the proposed research is in line with their expectations of health-related research and to discuss if they have any concerns regarding the research.

e. **Blood Donors Studies Steering Committee**. The Blood Donors Studies Steering Committee safeguards the wellbeing of the blood donors and monitors the overall conduct of the studies including aspects relating to the BioResource. The DAC may seek advice regarding an application from the Blood Donors Studies Steering Committee if required.

f. The steering committee membership includes an independent chairperson (Professor Jane Armitage, University of Oxford), senior NHSBT staff, independent trial experts, a sponsor representative and blood donor representatives.

4.3 Registration

a. This stage is to confirm the identity of the applicant (and any co-applicants) seeking to use the BioResource and to check their *bona fides* before adding them to the database of applicants.

b. Researchers requiring data from the BDSB will email the BTRU Research Programme Manager to request a Project Proposal form.

c. At this point the applicant is added to a database in order to track all aspects of the request process. All communications relating to each application will be retained so that there is a comprehensive file.

d. Each applicant wishing to use the BioResource must provide contact details and employment information (research department, address, email and website) for themselves and any co-applicants.

e. The Research Programme Manager will take the steps needed to confirm the identity and the *bona fides* of the applicant(s). This may include confirming they hold an email account at their host institution or otherwise validating their employment and checking that they and their employing organisation have the capability to carry out the planned research, for example by reviewing their publication history. When this has been done satisfactorily, the applicant will be issued with a project number for the request.

4.4 Preliminary Application

a. The applicant must provide, via email, brief information about the project and data required, to enable the Research Programme Manager to determine whether the application should progress.

b. The Research Programme Manager will review the information provided as part of the Preliminary Application and seek any further information that is required (which may relate to the scientific basis, feasibility or some other characteristic of the proposed research project).

c. The Research Programme Manager will also determine whether the BioResource contains the data required for the proposed research.

d. When the review has been completed, the Research Programme Manager will communicate one of the following responses to the applicant:
   - An invitation to complete the main application by completing a Project Proposal Form. The Research Programme Manager will also endeavour to answer any of the applicant's questions; or
   - An indication that the application will likely be denied (with summary reasons given), which will typically be determined by the BDSB DAC. The applicant can then elect to (i) amend the application; (ii) withdraw the application; or (iii) request that the decision be reconsidered (in which case the process in section [4.7] will be followed).

4.5 Main Application

If the applicant has received an invitation to complete a Main Application, then they will need to complete a Project Proposal Form and provide the following information:

- Principal applicant contact details
- Co-applicant contact details (who also need to be registered as detailed in section [4.3])
- Name and email address of individual authorised to sign the appropriate agreement on behalf of the applicant's employer
- Project title
- Proposed timetable
- Lay summary
- Scientific rationale of project (background and any pilot data; experimental details and design; power calculations; expected value of results; relevant references)
- Rationale for using the BDSB, including how the research project supports the BioResource's purpose (health-related and public interest as per section [2.0b] of the Data Access Policy)
- Required data
- Date and signature

4.6 Response to Main Application Form

a. In the first instance, the Research Programme Manager will review the Project Proposal Form and seek any further information that is required (which may relate to the scientific basis, feasibility or some other characteristic of the proposed research project).

b. The Project Proposal Form will be made available to the DAC for review. Any questions or concerns raised by the DAC will be communicated to the applicant, via the Research Programme Manager.

c. The DAC will evaluate the proposal form to:

I. assess the applicant/co-applicants and ensure they are *bona fide* researchers with the appropriate experience to carry out the planned study based on, e.g. publication track record
II. review the specific project aims and justification and assess whether the data requested is relevant to the project aims
III. determine whether the proposed research use meets the required criteria for access to the data within the database (i.e. it represents high quality research into blood donation or wider health or health-related questions)
IV. ensure that data access requests do not carry a significant potential for participant identification (e.g. through extremely specific data requests)
V. ensure research projects have the relevant scientific approval if required.

d. When the review has been completed, the Research Programme Manager will communicate the DAC's decision to the applicant.

e. The DAC will make a decision to either:
- Approve the application subject to the applicant entering into an appropriate agreement
- Approve the application with conditions on outstanding matters;
- Deny the request for data access. The applicant can then elect to (i) amend the application; (ii) withdraw the application; or (iii) request that the decision be reconsidered.

f. The DAC may seek advice regarding an application from the Blood Donors Studies Steering Committee.

g. Applications for certain datasets may also need to be reviewed by other committees. For example applications that include INTERVAL whole genome sequencing (WGS) data will need approval from the INTERVAL WGS Steering Committee before the sequencing data is released.

4.7 Reconsideration of Applications

a. If an applicant is advised that the DAC is likely to decline a Preliminary Application or if the DAC decline the Main Application, then the applicant may request that the application be reconsidered by the DAC.

b. The process for having an application reconsidered is as follows:
- Within 1 month of the relevant decision, the applicant should submit a written request, giving their reasons why they consider that the decision should be revised.
- Within 4-6 weeks of receipt of such a request, the DAC will aim to consider it, along with the original application (and any other information that it considers pertinent), and the Research Programme Manager will then respond to the applicant.

- The DAC may seek additional advice from the BDS Steering Committee, in which case the applicant will be advised by the Research Programme Manager of any revision to the timetable for review.

c. If, following reconsideration, the application is declined, then the applicant will not be able to submit the same proposal again within a 12 month period (unless the BDSB DAC has indicated specifically that it may be submitted subject to specific changes being made).

4.8 Agreements

a. If the application is approved, then the University of Cambridge will prepare an appropriate agreement and provide this to the applicant for review and completion by their institution.

b. A Data Transfer Agreement or Research Collaboration Agreement may be required, depending on the nature of the application.

c. If an applicant breaches the provisions of the agreement then this could lead to immediate revocation of the licence to use the BioResource. It could also lead to other actions, such as informing the applicant's employer, as well as other regulatory bodies, and prohibiting further access to the BioResource from the applicant's institution.

4.9 Timelines

a. In order to facilitate access to the data in the BioResource, it is intended that each review stage will be conducted in accordance with an indicative timeline (although the need to seek further information or guidance on particular applications may lead to a more prolonged process).

b. Due to limited resources priority will be given to applicants whose work is closely aligned to the primary aims of the studies (i.e. enhancing the safety of blood donation and ensuring sustainability of the blood supply) and to strategic collaborations involving the study principal investigators.

c. The intended timelines for each stage are:
   - Registration: 1 week to check identity and add to database
   - Preliminary application: 3 weeks to review and respond
   - Main application: 3 weeks to review and respond
   - Agreements: Timelines are variable and dependant on other departments within the University of Cambridge and the applicant's institution.
   - Release of data: Data will be released as soon as the necessary agreement has been executed.

d. The process from the start of the review of the Preliminary Application to a decision takes approximately two months. For researchers outside the University of Cambridge, a formal agreement must then be put in place with the applicant's institution before data can be released.

5. Provision of Data

5.1 Data Format

a. Data from the BioResource will be provided in the following formats depending on the nature of the dataset:
   - Certain datasets will be made available by issuing hyperlinks and passwords to the applicant so that the relevant data can be accessed and downloaded. The hyperlinks and password are sent in separate emails and the password is only sent upon confirmation of receipt of the hyperlink email. Only applicant(s) are given access to the dataset; or
   - Certain datasets will be made available by sending a .zip file and passwords. The .zip file and password are sent in separate emails and the password is only sent upon confirmation of receipt of the email containing the .zip file; or
   - Large datasets will be transferred via Secure File Transfer Protocol (SFTP) or in situ access may be made available to the applicant on a University secure server.

b. The dataset provided is de-personalised and will contain only the subset of participants and data items required for the project from the de-personalised study database.

c. To further protect participant confidentiality, exact dates are perturbed by being randomly adjusted by a few days (+7/-7) for each individual or are replaced with month and year. Study identifiers are replaced with a new identifier unique to that dataset.

d. Each dataset is assigned unique mapping IDs so that applicants cannot combine datasets.

5.2 Applicant Data Security

a. Data should only be accessible to the Applicant(s) named on the research proposal.

b. Files should either have applicant-only read/write access (not group or world access) or project-specific groups should be used for group access that contains only those names authorised to access the data.

c. Data kept on laptops should be encrypted when not in active use, either in individual encrypted files or in encrypted directories/partitions.

d. Data must not be held on USB keys or other portable hard drives.

5.3 Terms and Conditions

a. Applicants shall handle all data in accordance with the General Data Protection Regulation (GDPR) 2016/679 and/ or any applicable local legislation.

b. Safeguards must be maintained to help ensure the anonymity and confidentiality of participants' data. The Applicant(s) must not make any attempt to identify participants.

c. If the Applicant(s) is part of two or more data access agreements, the Applicant(s) must not merge data together across projects.

d. In the case of inadvertent identification, researchers are required to report this immediately to the BDSB and not to disclose it further or make any attempt to contact the individual.

e. The applicant is required to destroy all copies of the provided data or otherwise render it inaccessible once the project is complete.

6. Publication of Findings and Return of Results

a. The applicant is required to use their best endeavours to publish the findings of any research deriving from the data in the BioResource in an academic journal or on an open source publication site within the amount of time stated in the agreement.

b. It is mandatory for the Applicant(s) to acknowledge data from the BioResource and its funders in any publications arising from the research project. Applicants must contact the Research Programme Manager to ensure that the correct text is used in publications. Representatives of the studies should be offered co-authorship.

c. Any new data or derivations (e.g. from new assays, e-health linkages, summary statistics, database cross-referencing) created through the research project must be fully provided within three months of generation to be merged into the BioResource.

d. The BDSB will give reasonable consideration to written requests (containing an appropriate explanation) for an extension of these time limits.

7. Recording and communicating use of the data

7.1 Recording details of applicants and applications

a. The Research Programme Manager will maintain a database of applications to use the BDSB. This will include details of:
   - Project title
   - Project summary
   - The dataset requested
   - Details of the principal applicant, their institution and contact details
   - Details of any co-applicants
   - The location of the research
   - Whether the application was approved
   - Start and end dates (if approved)

b. Together with any other details necessary to manage the application progress and track the progress of approved projects.

7.2 Communicating the use of data

a. The BDSB is committed to maintaining a dialogue with blood donors, researchers and the public at large and will keep these communities updated on the progress of the BioResource and the research work which is carried out using the data.

b. Summary details of approved applications will be published on the Blood and Transplant Research Unit in Donor Health and Genomics' website

(www.donorhealth-btru.nihr.ac.uk). These will include the project title, the name and institution of the principal investigator and the lay summary.

c. Summaries and/or links to publications that derive from use of the data will be provided on the study websites; INTERVAL (www.intervalstudy.org.uk), COMPARE (www.comparestudy.org.uk) and STRIDES (www.strides-study.org.uk). Updates will also be available on the BTRU website.

d. Research findings will be highlighted in study newsletters, which are sent to blood donors who participated in a blood donors study and consented to receiving newsletters.

e. Contentious and/or ethically challenging issues related to proposed or approved uses of the BioResource will be highlighted on the website. This will allow blood donors, and the wider public, to provide input on particular research uses and other issues.

Appendices

Appendix 1- INTERVAL Study data

The University of Cambridge currently holds the following information for the INTERVAL participants:

- Questionnaire information
  - Epidemiology questionnaire data at 0, 6, 12, 18, 24, 30, 36, 42 and 48 months.
  - Quality of life survey data at 0, 6, 12, 18, 24, 30, 36, 42 and 48 months.

- Blood donation records

  Information retrieved from NHSBT's national database (PULSE), including data on donors' ethnicity, blood group, NHSBT registration date, date of first attendance for donation, date of first successful donation and detailed information on previous donations and deferral history.

- Donation deferrals.

- Adverse events occurring during or after the donation visit.

- Physical activity
  7-day accelerometry data from a subset of randomly selected participants and online physical activity questionnaire at the same timepoint.

- Cognitive function
  Online cognitive function tests, including Stroop Test (attention and reaction times), Trail Making Test (executive function), Pairs Test (Episodic Memory) and Reasoning Tests (intelligence).

- Full blood count data and extended parameters for the 0 and 24 months timepoints.

- Molecular assays
  Data from several molecular assays have been recorded either on a cohort-wide basis or in large subsets:
  - Basic clinical chemistry assays including ferritin, iron, transferrin, transferrin saturation, total iron binding capacity, hepcidin, C-reactive protein, fructosamine, HbA1c, triglycerides and lipids;
  - Foetal haemoglobin;
  - Haematological parameters, including ~80 blood cell traits;
  - Plasma proteomics, including >4000 proteins measured with different technologies (e.g., aptamers [SomaLogic] and antibodies [Olink]);
  - Plasma metabolites, including >1100 metabolites measured using untargeted mass-spectrometry;
  - Plasma lipids, including >1200 lipids measured using complementary NMR and high-resolution MS-based assays.

  - Genotyping array (~830K measured variants including GWAS backbone & Exome array content) assayed and then imputed to 1000G+UK10K combined reference panel (80M variants in total);
  - Whole genome/exome sequencing (15X WGS; 50X WES).

- Data from Electronic health records
  - Hospital Episode Statistics obtained from NHS Digital
  - Cancer and death registrations obtained from NHS Digital

Appendix 2- COMPARE Study data

The University of Cambridge currently holds the following information for the COMPARE participants:

- Questionnaire information
  - Epidemiology questionnaire
  - Quality of life survey data

- Blood donation records
  Information retrieved from NHSBT's national database (PULSE), including detailed information on previous donations and deferral history.

- Donation deferrals

- Adverse events occurring during or after the donation visit.

- Full blood count data and extended parameters.

- Molecular assays
  Data from several molecular assays have been recorded either on a cohort-wide basis or in large subsets:
  - Genotyping

- Data from Electronic health records
  - Hospital Episode Statistics obtained from NHS Digital
  - Cancer and death registrations obtained from NHS Digital

Appendix 3- STRIDES BioResource Data

The University of Cambridge is collecting the following information for the STRIDES BioResource participants:

- Questionnaire information
  - Epidemiology questionnaire
  - Quality of life survey data
  - Follow- up questionnaire with questions relating to iron deficiency and restless leg syndrome.

- Blood donation records

Information retrieved from NHSBT's national database (PULSE), including detailed information on previous donations and history of vasovagal reactions.

- Donation deferrals
- Adverse events occurring during or after the donation visit.
- Full blood count data and extended parameters.
- Data from molecular assays
- Data from electronic health records

Appendix 4- Related Documents

- Project Proposal Form
- Data dictionaries
- Study Consent forms/ Information leaflets